

CLAIMS

- 1 1. A predictive method for admitting a content request to an
2 information system, comprising:
 - 3 receiving a user's quality of service objectives at the information
4 system;
 - 5 receiving a content request to be admitted to the information system;
 - 6 determining if the content request is for an existing session or a new
7 session; and
 - 8 sending the content request to a dispatch control function at the
9 information system when the content request is for an existing session.
- 1 2. The method of claim 1, wherein the user's quality of service
2 objectives include speed of content delivery for a specified time.
- 1 3. The method of claim 1, wherein the user's quality of service
2 objectives include consistency of speed of content delivery.
- 1 4. The method of claim 1, wherein the user's quality of service
2 objectives include a function of number of concurrent users.
- 1 5. The method of claim 1, wherein the user's quality of service
2 objectives include system response time.
- 1 6. The method of claim 1, wherein the user's quality of service
2 objectives include system response time consistency.

- 1 A reactive method for admitting a content request to an
2 information system with a server, comprising:

3 receiving a user's quality of service objectives at the information
4 system;
5 receiving a content request to be admitted to the information system;
6 calculating a quota of maximum sessions that the server can handle
7 and maintain the user's quality of service objectives; and
8 determining if the content request exceeds the quota.

1 8. The method of claim 7, further comprising:
2 sending the content request to the server if the quota is not exceeded.

1 9. The method of claim 7, further comprising:
2 if the quota is exceeded and the user's quality of service objectives
3 are met, then sending the content request to a throughput computation to
4 determine whether or not the server can process any more content requests.

1 10. The method of claim 7, further comprising:
2 if the quota is exceeded and the user's quality of service objectives
3 are not met, then rejecting the content request.

1 11. The method of claim 7, further comprising:
2 if the quota is exceeded and the user's quality of service objectives
3 are not met, then downgrading the user's quality of service objectives and
4 sending the content request to a throughput function whether or not the
5 server can process any more content requests.

1 12. The method of claim 9, wherein the throughput calculation is
2 a capacity utilization of the server using content request arrival rates,
3 latencies and a quota of maximum content requests that a server can handle
4 and maintain the user's quality of service objectives.

1 13. The method of claim 7, wherein the quota calculation is done
2 by observing a fixed number of content requests.

1 14. The method of claim 7, wherein the quota calculation is
2 determined by the number of times the user's quality of service objectives
3 have been violated divided by the number of content requests.

1 15. A method for admitting a content request to an information
2 system, comprising:

3 receiving a user's quality of service objectives at the information
4 system;

5 receiving a content request to be admitted to the information system
6 from a requestor;

7 determining if the content request is for an existing session or a new
8 session; and

9 if the content request is not part of an existing session then
10 predicting future content requests expected in a predetermined time for the
11 information system.

1 16. The method of claim 15, further comprising:

2 aggregating new content requests expected in the predetermined
3 time with existing content requests currently being processed by the
4 information system and create an aggregated content capacity request.

1 17. The method of claim 16, further comprising:

2 determining if the information system can process the aggregated
3 content capacity request in compliance with the user's quality of service
4 objectives.

1 18. The method of claim 17, further comprising:

2 accepting or rejecting the content request.

1 19. The method of claim 18, further comprising:
2 sending the content request to dispatch control if the content request
3 is accepted.

1 20. The method of claim 18, further comprising:
2 sending the content request to a user defined rejection rule if the
3 content request is rejected.

1 21. The method of claim 20, wherein the user defined rejection
2 rule includes sending a message to the requestor.

1 22. The method of claim 20, wherein the user defined rejection
2 rule includes making the information system unavailable for the content
3 request.

1 23. The method of claim 20, wherein the user defined rejection
2 rule includes making the information system unavailable for the content
3 request for a selected period of time.

1 24. The method of claim 20, wherein the user defined rejection
2 rule includes queuing the content request for admission to the information
3 system.

1 25. The method of claim 20, wherein the user defined rejection
2 rule includes gracefully degrading a quality service compliance of sessions
3 currently existing in the information system.

1 26. The method of claim 20, wherein the user defined rejection
2 rule includes gracefully degrading a quality service compliance of new
3 sessions incoming to the information system.

1 27. The method of claim 20, wherein the user defined rejection
2 rule includes gracefully degrading new and existing sessions.

1 28. The method of claim 20, wherein the user defined rejection
2 rule includes gracefully degrading lower priority customers as defined in
3 users' quality of service objectives.